Journal of Chemical Ecology, Vol. 30, No. 8, August 2004 (© 2004)

STATISTICAL PROBLEMS ENCOUNTERED IN TRAPPING STUDIES OF SCOLYTIDS AND ASSOCIATED INSECTS

JOHN D. REEVE[1,*] and BRIAN L. STROM[2]

[1]*Department of Zoology*
Southern Illinois University
Carbondale, Illinois 62901, USA

[2]*Southern Research Station*
USDA Forest Service
Pineville, Louisiana 71360, USA

(Received September 8, 2003; accepted April 2, 2004)

Abstract—Traps baited with semiochemicals are often used to investigate the chemical ecology of scolytids and associated insects. One statistical problem frequently encountered in these studies are treatments that catch no insects and, thus, have zero mean and variance, such as blank or control traps. A second problem is the use of multiple comparison procedures that do not control the experimentwise error rate. We conducted a literature survey to determine the frequency of these two statistical problems in *Journal of Chemical Ecology* for 1990–2002. Simulations were then used to examine the effects of these problems on the validity of multiple comparison procedures. Our results indicate that both statistical problems are common in the literature, and when combined can significantly inflate both the experimentwise and per comparison error rate for multiple comparison procedures. A possible solution to this problem is presented that involves confidence intervals for the treatment means. Options to increase the statistical power of trapping studies are also discussed.

Key Words—Scolytidae, multiple comparisons, experimentwise error rate, blank treatment, homogeneity of variances, semiochemical, pheromone.

INTRODUCTION

Bark beetles (Scolytidae) and their associates illustrate many important phenomena in chemical ecology. Studies involving these organisms have elucidated the role of aggregation pheromones and host volatiles in the colonization process of the host tree, the role of antiaggregation pheromones that apparently limit

*To whom correspondence should be addressed. E-mail: jreeve@zoology.siu.edu

attack densities, and the use of kairomones by natural enemies and competitors to locate prey and resources (Borden, 1982; Wood, 1982; Smith et al., 1993; Raffa, 2001). Novel methods of control have also been developed that use these chemical signals to deflect beetle attack from host trees (Borden, 1997). Trapping experiments in the field are frequently used to study the chemical ecology of these organisms. These studies are often designed to address two basic questions: (1) which treatments (chemicals or biological material) are attractive to the insects, and (2) which treatments increase or decrease catches of insects relative to other treatments? Both questions are typically examined within the same trapping study. Completely randomized or randomized block designs are commonly used, and the resulting data are counts of the number of insects caught in each trap.

One statistical problem frequently encountered in semiochemical trapping studies are unequal variances (heteroscedasticity) among treatments, with variance-stabilizing transformations such as $\sqrt{Y}$or $\log(Y + 1)$ typically used as a remedy. In experiments that involve blank traps or other treatments that are unattractive, however, it is often the case that all the observations are zero and so have zero variance. Variance-stabilizing transformations are not useful here, because they cannot create variance where none exists. Treatments that have zero or low variance will reduce the magnitude of mean square error, which is the denominator for F tests and also a component of multiple comparison procedures used to compare treatments. Thus, blank traps and other low variance treatments could potentially affect the statistical analysis of semiochemical trapping studies by increasing apparent treatment effects.

Another problematic feature of trapping studies is the use of multiple comparison procedures that do not control the experimentwise error rate. The experimentwise error rate is defined as the probability of one or more Type I errors (spurious results) in a set of comparisons, usually all pairwise comparisons among treatments in an experiment. One commonly used procedure is Fisher's protected LSD (least significant difference) test, which is well-known to have an experimentwise error rate that increases rapidly with the number of treatments in the experiment (Hayter, 1986; Toothaker, 1993; Hsu, 1996; Westfall et al., 1999). Thus, as the number of treatments increases the probability of finding at least one significant difference by chance also increases. Some authors have argued that it is more appropriate to control the per comparison error rate (Carmer and Swanson, 1973; Carmer and Walker, 1982; Rothman, 1990; Saville, 1990; Stewart-Oaten, 1995), defined in this context as the probability of Type I error for a single pairwise comparison between treatments. Regardless of the merit of these arguments, readers of trapping studies should be aware that the two types of multiple comparison procedure have different statistical goals and are not equivalent. Methods that control the experimentwise error rate place a premium on controlling all Type I errors and so are inherently more conservative procedures. The problem with this approach is that differences that do exist may not be detected because of low

statistical power. Methods outside this category (such as Fisher's protected LSD) are more powerful but are also more likely to find spurious differences among treatments.

We had three objectives in this study. The first was to document the prevalence of the statistical problems discussed above in published studies of scolytids and their associated insects (typically competitors and predators). We conducted a literature survey to determine the frequency of blank treatments in trapping studies, the number of treatments per experiment, and the multiple comparison procedure used. The multiple comparison procedures were also classified into two groups depending on whether they control the experimentwise error rate. We confined our survey to papers on these organisms because of our familiarity with their objectives and to reduce the papers surveyed to a manageable number. The second objective was to examine the effect of blank treatments on the validity of multiple comparison procedures, in particular the experimentwise and per comparison error rates. We used simulation studies to evaluate these rates for two disparate procedures, one that controls the experimentwise error rate (Tukey's HSD or honestly significant difference) and one that does not (Fisher's protected LSD). Our third objective was to develop an alternate method of analysis that avoids these statistical problems. Various options to increase the statistical power of semiochemical trapping studies are also discussed.

METHODS AND MATERIALS

Literature Survey. We surveyed papers published in the *Journal of Chemical Ecology* involving trapping studies of scolytids and associated insects for the interval 1990–2002. For each paper, we determined whether any experiments incorporated blank traps, the average number of treatments per experiment (most papers involved several experiments), whether the analysis was parametric or nonparametric, and if parametric the multiple comparison procedure used (if any). The multiple comparison procedures we encountered were Fisher's protected LSD, simple LSD with no preliminary ANOVA, Duncan's multiple range test, Student–Newman–Keuls multiple range test, the REGW multiple range procedures in SAS (SAS Institute Inc., 2001), Tukey's HSD, Dunnett's test (which compares treatments with a control), and a Šidák adjustment for multiplicity. General descriptions of these procedures can be found in Sokal and Rohlf (1995), Hsu (1996), and Westfall et al. (1999). Of these procedures, only the last four control the experimentwise error rate (Day and Quinn, 1989; Hsu, 1996; Westfall et al., 1999).

Simulation Studies. We conducted simulations to estimate the experimentwise error rate for two different multiple comparison procedures, Fisher's protected LSD and Tukey's HSD, in randomized block experiments involving a blank trap treatment. These procedures were chosen because they span the range of control

of the experimentwise error rate (none to strong) and are easy to simulate. Fisher's protected LSD is also one of the most commonly used multiple comparison procedures in the literature. Blank traps were assumed to catch zero insects and thus have zero variance, as occurs in many pheromone trapping studies (e.g., Herms et al., 1991; Miller et al., 1997; Zhou et al., 2001). All other treatments were defined to have the same mean and variance, implying there are no real treatment effects in the experiment beyond the blank treatment. We were interested in determining the error rate for all pairwise comparisons among these other treatments, excluding any comparisons involving blank traps versus other treatments. These were excluded because the null hypothesis was always false in this case (blank traps by assumption catch fewer insects than any other treatment). We also examined how the experimentwise error rate varied with the number of treatments in the simulations.

The parameter values in the simulations were based on trapping experiments (seven total) involving the scolytid predator *Thanasimus dubius* (Coleoptera: Cleridae) (J. D. Reeve, B. L. Strom, L. Rieske-Kinney, and B. D. Ayres, unpublished data). The treatments involved various bark beetle pheromones and tree volatiles, and typically captured 10–100 adult predators per trap during the course of the experiment, except for blank traps that caught virtually no insects. The data were transformed before analysis using the log transformation $\log_{10}(Y + 1)$. On this scale of measurement, we observed a mean square error $\sigma^2 \approx 0.06$, after excluding unattractive treatments from the analysis. The variance due to the block effect was considerably smaller than σ^2, and based on these studies we chose two different values for the simulations, $\sigma_B^2 = 0$ or 0.03. In terms of standard deviations, these values correspond to $\sigma = 0.245$ and $\sigma_B = 0$ or 0.173. Observations for the simulations were generated using a standard statistical model for randomized block designs:

$$Y_{ij} = \mu + \alpha_i + B_j + \varepsilon_{ij},$$

where μ is the grand mean, α_i is the effect of ith treatment, B_j is a normal random variable with mean zero and variance σ_B^2, and ε_{ij} is normal with mean zero and variance σ^2. This model specifies that treatments are fixed while blocks are random effects. Block was considered a random effect because the blocks used in semiochemical studies are typically a sample of possible study sites. We used $\mu = 1.1$ as the grand mean in the simulations, corresponding to the average number of insects captured in our experiments (on a log scale). Treatment effects were assumed to be absent, implying that $\alpha_i = 0$ for all i. Observations for the blank treatment were generated by forcing Y_{ij} to zero for that treatment across all blocks. A total of 10 blocks was used in each replicate simulation. We also varied the number of treatments in the experiment, ranging from 3 to 10 treatments. In a second set of simulations, we added another unattractive treatment to the experiment, so that there were two with zero mean and variance (a fairly common

occurrence in trapping studies). We note that the F test for the treatment effect was always significant using these parameter values, and the block effect frequently so when $\sigma_B^2 = 0.03$. The simulations were programmed using the statistical language R 1.7.0 (The R Development Core Team, 2003).

Each simulated data set was analyzed using ANOVA for randomized block designs followed by multiple comparisons using Fisher's protected LSD or Tukey's HSD. For Fisher's protected LSD, we used $\alpha = 0.05$ as the significance level for both the ANOVA and pairwise comparisons, whereas for Tukey's HSD we used an experimentwise error rate of 0.05. For each data set, the program counted the number of significant pairwise comparisons, excluding those involving the blank traps. The experimentwise error rate was estimated as the fraction of data sets having at least one significant comparison (a Type I error) in 5000 replicate simulations. We also computed the experimentwise error rate assuming the blank treatment had the same variance as other treatments, for comparison with the rates obtained with zero variance (the simulations). Here, we were able to directly calculate the error rate using the Studentized Range distribution, and in the case of Fisher's protected LSD, by assuming that the treatment F test was significant, as would be expected given the hypothesized treatment means (Hayter, 1986; Hsu, 1996). This error rate can be regarded as a predicted error rate for Fisher's protected LSD and Tukey's HSD without the zero variance problem caused by blank traps.

Using the same simulations, we also estimated the per comparison error rate for Fisher's protected LSD. This was estimated as the fraction of data sets having a significant comparison between an arbitrary pair of treatments, again excluding the blank treatment from consideration.

RESULTS

Literature Survey. The survey indicates that experiments involving blank treatments are common in pheromone trapping studies of scolytids and associates. Of the 50 papers that met our criteria, 84% (42 of 50) of the papers contained a blank or control treatment, which typically caught few or no insects. Parametric procedures were used in 78% of the papers (39 of 50), with 92% of these (36 of 39) using various multiple comparison procedures to test for differences among treatments. Averaging across all papers, the mean number of treatments per experiment was 5.66 (SD = 3.03, range 2–20). Of the papers using multiple comparison procedures, 56% (20 of 36) used methods that do not control the experimentwise error rate (Table 1). The most commonly used methods were LSD (either Fisher's protected LSD or simple LSD) and the REGW procedures of SAS (SAS Institute Inc., 2001). Of those papers using multiple comparisons, 42% (15 of 36) used methods that failed to control the experimentwise error rate and also had blank treatments, meaning they had both of these statistical problems.

TABLE 1. FREQUENCY OF VARIOUS MULTIPLE COMPARISON PROCEDURES IN JOURNAL OF CHEMICAL ECOLOGY PAPERS FOR 1990–2002

Procedure	Number of papers	Control experimentwise error rate?
Duncan's multiple range	6	No
LSD	10	No
Student–Newman–Keuls	4	No
Dunnett's test	1	Yes
REGW	12	Yes
Šidák	1	Yes
Tukey's HSD	2	Yes
Total	36	

Note. Procedures were classified as controlling the experimentwise error rate based on Hsu (1996) and Westfall et al. (1999).

Simulation Studies. The experimentwise error rates for Fisher's protected LSD and Tukey's HSD are shown in Figure 1, for simulations involving a single blank treatment with zero mean and variance. The experimentwise error rate for the protected LSD increases rapidly with the number of treatments and a blank treatment increases the rate even further, as compared to the predicted error rate for this procedure (Figure 1A). However, the effect of the blank treatment was less in the presence of a block effect ($\sigma_B^2 = 0.03$). This likely occurs because the blank treatment generates a Treatment × Block interaction in this situation (see Discussion). The experimentwise error rate for Tukey's HSD is not influenced by the number of treatments, as would be expected, and is somewhat elevated by the blank treatment although the overall rate remains low. As with the protected LSD, a block effect reduces the influence of the blank treatment on the experimentwise error rate. The addition of a second unattractive treatment increases the experimentwise error rate even further, especially in comparison to the predicted error rate for each procedure (Figure 2).

We also examined the per comparison error rate for Fisher's protected LSD, because this is the error rate nominally controlled by this procedure. The simulations indicate that the per comparison rate is elevated by a blank treatment above the 0.05 level, especially if there is a second unattractive treatment in the experiment (Figure 3). The effect on the per comparison rate diminishes as the number of treatments increases, presumably because a blank treatment has less effect on mean square error when there are more treatments. A strong block effect ($\sigma_B^2 = 0.03$) also reduces the effect of the blank treatment on the per comparison error rate.

Additional simulations (not shown) using different values of σ^2 and the number of blocks yielded virtually identical results to those in Figures 1–3, suggesting these patterns are independent of the actual parameter values chosen.

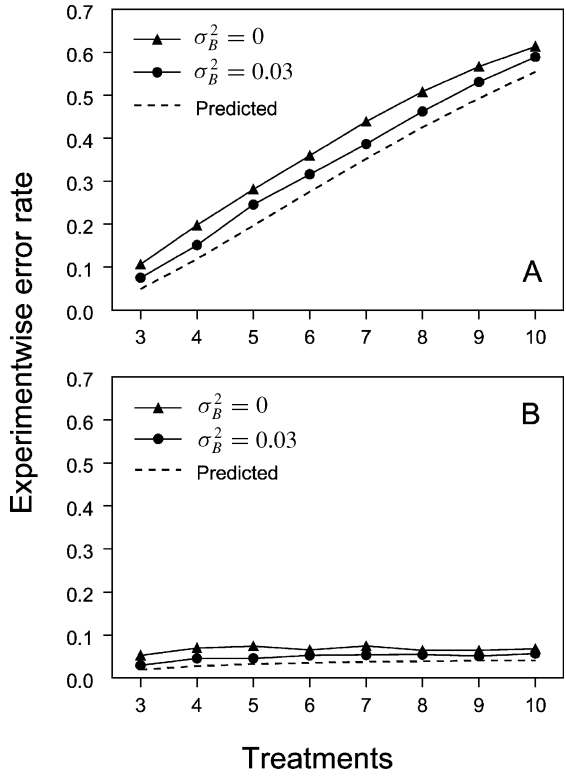

FIG. 1. Experimentwise error rate for Fisher's protected LSD (A) and Tukey's HSD (B) in simulated experiments incorporating a blank treatment with zero mean and variance, and no other treatment effects. Error rates are plotted as a function of the number of treatments and the presence or absence of a block effect ($\sigma_B^2 = 0$ vs. 0.03). The dashed line is the predicted error rate when the blank treatment has the same variance as other treatments, calculated using the Studentized Range distribution.

DISCUSSION

The literature survey indicates that blank, unattractive treatments are a common feature of trapping studies, as are multiple comparison procedures that do not control the experimentwise error rate. The simulations involving Fisher's protected LSD suggest the combination of these two features can lead to high experimentwise error rates. It can also substantially elevate the per comparison error rate, the only rate really controlled by this procedure. Fisher's protected LSD thus seems less than ideal for analyzing data of this type, especially when blank traps are present. The experimentwise error rate for Tukey's HSD was also increased by

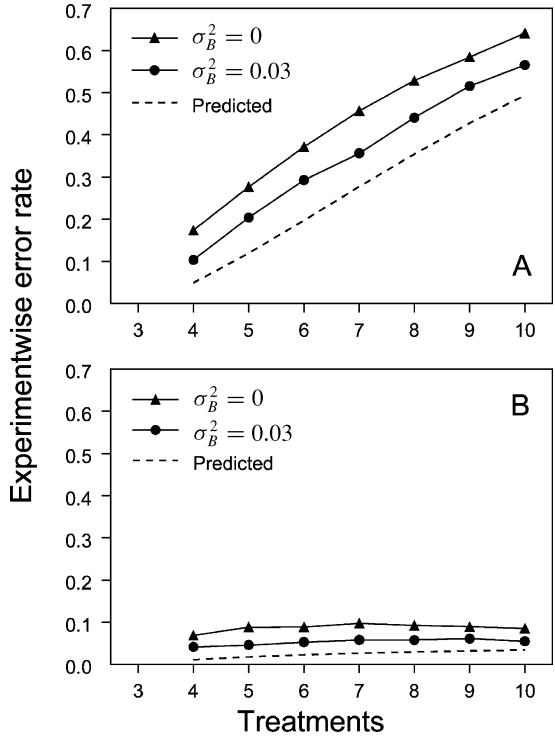

FIG. 2. Experimentwise error rate for Fisher's protected LSD (A) and Tukey's HSD (B) in simulated experiments incorporating two treatments with zero mean and variance, and no other treatment effects. For further details see Fig. 1.

unattractive treatments, but was not a function of the number of treatments and remained relatively low in any event. This result suggests that authors interested in controlling the experimentwise error rate could use a procedure such as Tukey's HSD even with blank treatments in the experiment. The mechanism underlying these results is the reduction in mean square error caused by zero variance treatments in the data, causing it to underestimate the variability among observations in the other treatments in the design. Many other multiple comparison procedures use mean square error in their formulation and we would expect their error rates to also be inflated, although the rates for procedures that control the experimentwise error rate should be lower.

An interesting finding in the simulations was the interplay between the block effect and the per comparison and experimentwise error rates. The presence of a block effect reduces these error rates apparently because it generates a Treatment $\times$ Block interaction in the data. The source of this interaction is the fact that blank

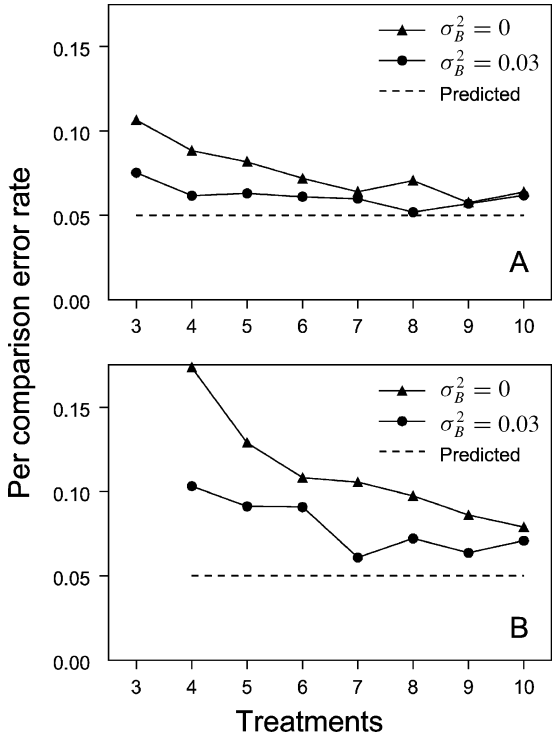

FIG. 3. Per comparison error rate for Fisher's protected LSD in simulated experiments incorporating one (A) or two (B) treatments with zero mean and variance, and no other treatment effects. The dashed line is the predicted per comparison error rate when the blank treatment has the same variance as other treatments.

treatments always have zero mean regardless of any differences in block means. The statistical model for randomized block designs has no interaction term to absorb this source of variation, so instead it inflates the value of mean square error, partially countering the deflating effect of the blank treatment. This result implies that the per comparison and experimentwise error rates are to some extent indeterminate when blank treatments are included in the analysis, because they apparently depend on the magnitude of the block effect.

A possible solution to these statistical problems would be to remove the blank treatment (and possibly other zero variance treatments) from the analysis. This would ensure that error rates are near their specified values, whatever multiple comparison procedure is used. The elimination of unattractive treatments from the data would also facilitate analysis using generalized linear mixed models, an extension of mixed models to Poisson or other discrete distributions (McCulloch

and Searle, 2001). In our experience, the software developed for these models can have numerical problems with concentrations of zeros.

Removal of the blank treatment does present a problem in evaluating which treatments are attractive to insects (a common objective of semiochemical trapping studies) because it leaves no standard with which to compare the remaining treatments. One way of dealing with this problem is to specify some minimum level of attraction believed to have biological relevance and compare the treatments to this level. For example, an investigator could specify that a treatment must exceed an average of one insect per trap during the course of the experiment to be considered attractive. Confidence intervals for the treatment means could be used to make this decision more quantitative and are more informative for this reason than the usual standard errors. If we are only interested in treatments that exceed a specified level, then one-sided confidence intervals may be appropriate. If the confidence interval boundary exceeds the specified level of attraction, this actually constitutes an α-level test of H_0: $\mu = \mu_0$ versus A: $\mu > \mu_0$, where μ_0 is the specified level of attraction and $100(1 - \alpha)\%$ is the confidence level of the interval. We also suggest applying a Bonferroni correction to the value of α for the confidence intervals because we are constructing multiple intervals (one for each treatment). Finally, if transformations are used in the analysis it may be necessary to back-transform the confidence intervals to the original scale of measurement for comparison with μ_0 (or transform μ_0 itself). A sample SAS program and data set that illustrates these calculations is given in Appendix A. The sample data are trap catches of *T. dubius* from a randomized block experiment with six blocks and five treatments, including a blank treatment that caught no insects.

The program in Appendix A uses a standard mixed model for randomized block designs that assumes the data are normally distributed. Semiochemical trapping data are typically in the form of counts, however, and so this methodology may not be ideal, especially if a low number of insects are trapped. Generalized linear mixed models provide an alternative method of analysis for data of this type, and are implemented in SAS by the GLIMMIX macro (Littell et al., 1996; SAS Institute Inc., 2001). Appendix B lists a sample SAS program that uses this method to find confidence intervals for the treatment means. The model assumes the data are Poisson in distribution but also allows for over- or underdispersion in the observations. One and two-sided intervals for these data are shown in Figure 4, along with a reference line suitable for testing H_0: $\mu = 1$ versus A: $\mu > 1$, where μ is the mean number captured. We would reject this null hypothesis for any treatment whose one-sided lower confidence interval lies above this line, and accept it otherwise.

So far, we have dealt with problems of Type I error without addressing issues of statistical power. *Power* is defined as the probability of rejecting a null hypothesis when it is false and some alternative is true. Multiple comparison procedures involve the testing of multiple hypotheses, however, and several definitions of power have been developed for this situation (see Westfall et al.,

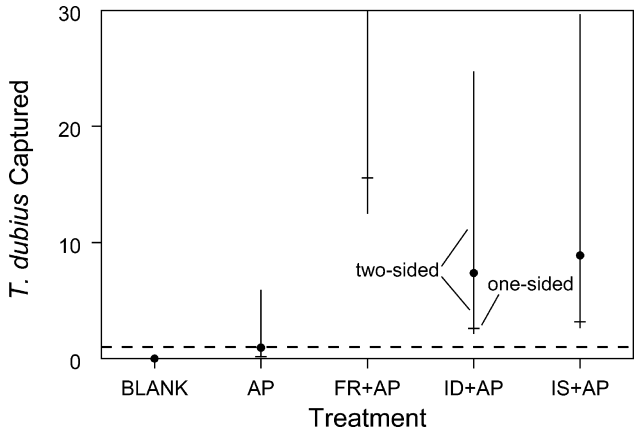

FIG. 4. One- and two-sided confidence intervals obtained using the sample data and program listed in Appendix B. The observations were obtained in a trapping experiment involving five semiochemical treatments (BLANK = blank trap, AP = α-pinene, FR + AP = frontalin + α-pinene, ID + AP = ipsdienol + α-pinene, IS + AP = ipsenol + α-pinene). The graph is truncated from above to show the confidence intervals for the less attractive treatments in more detail. A reference line for testing H_0: $\mu = 1$ vs. A: $\mu > 1$ is also shown.

1999). The easiest to calculate is individual power, defined as the probability of finding a significant difference between a given pair of treatments that differ by amount δ. We conducted a power analysis to examine how individual power varies with δ and the number of treatments and blocks in a randomized block design, using Tukey's HSD as the multiple comparison procedure. Power is also a function of mean square error or σ^2, and we used the same value as obtained in our trapping study of *T. dubius* ($\sigma^2 = 0.06$) after log transformation of the data (see Methods and Materials for further details). We used $\delta = 0.3$, 0.5, and 0.7 and five versus 10 treatments in the analysis. These values of δ correspond to approximately two-, three-, and five-fold differences in the treatment means on the original scale of measurement (before transformation). The results of the power analysis are shown in Figure 5. Power values of 0.8 or greater are typically regarded as sufficient (Cohen, 1988), and by this criterion we would have adequate power to detect three or five-fold differences between a pair of treatments using just eight blocks, although two-fold differences would require a much larger number of of blocks. This not a prohibitively large number of traps, and demonstrates that it is possible to have adequate power as well as control of the experimentwise error rate.

There are also ways of increasing power involving changes in only the experimental design, rather than increasing the number of traps deployed. If an experiment can be constructed as a comparison of a control treatment with other treatments, then one could use Dunnett's test as the multiple comparison procedure.

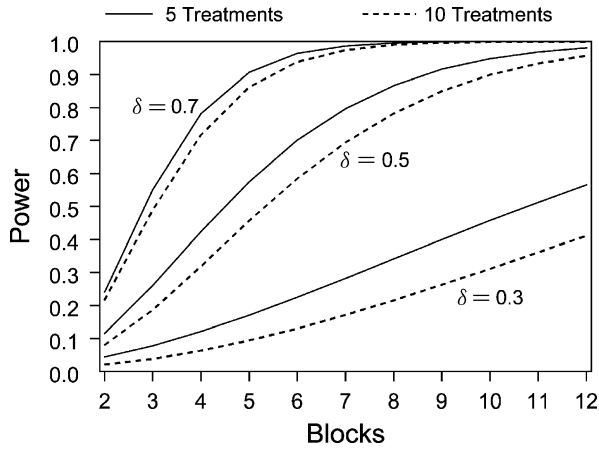

FIG. 5. An analysis of individual power for Tukey's HSD as a function of the number of blocks and treatments in the experimental design. Here δ is the difference between a given pair of treatments on a log (base 10) scale. See text for further details.

This procedure only involves pairwise comparisons of other treatments with the control, rather than all possible pairwise comparisons, and because it involves fewer comparisons is a more powerful test for experiments of this type (Hsu, 1996). We also suggest that investigators carefully consider the utility of including blank traps in every experiment for a particular insect species.

For experiments where the response to blank traps in unknown this treatment is a necessity. Blank traps will also be required in dose-response or other experiments in which it can reasonably be expected that some treatments may be negligibly attractive. In such experiments, blanks serve as a critical reference point, particularly under changing field conditions. However, in experiments comparing semiochemicals that are known to be attractive, blank traps may be wasteful particularly if previous studies have shown that they are unattractive. Resources devoted to implementing a blank treatment may be better utilized in increasing the sample size of other treatments, in particular the number of blocks for randomized block designs.

APPENDIX A

Sample SAS program and data set to find confidence intervals for the treatment means in a randomized block design using PROC MIXED in SAS (SAS Institute, 2001). The program first deletes the blank treatment observations, so that there are four remaining treatments in the analysis (see Discussion), and then log-transforms the observations. The confidence intervals for the treatment

means are generated by the lsmeans statement. To obtain a two-sided interval with a Bonferroni correction, we used $\alpha = 0.05/4 = 0.0125$ in the lsmeans statement. One-sided lower confidence intervals can be obtained by specifying twice this value ($\alpha = 0.025$) and using just the lower boundary. The intervals are then back-transformed to the original scale of measurement.

```
data traps;
      input block treat $ count;
      if treat="BLANK" then delete;
      logcount = log10(count+1);
      datalines;
1     AP      4
1     BLANK   0
1     FRAP   79
1     IDAP    7
1     ISAP   10
2     AP      1
2     BLANK   0
2     FRAP  124
2     IDAP   13
2     ISAP   20
3     AP      0
3     BLANK   0
3     FRAP   14
3     IDAP    .
3     ISAP    2
4     AP      0
4     BLANK   0
4     FRAP   15
4     IDAP   11
4     ISAP    7
5     AP      0
5     BLANK   0
5     FRAP   29
5     IDAP    7
5     ISAP    7
6     AP      2
6     BLANK   0
6     FRAP   70
6     IDAP   14
6     ISAP   20
;
run;
```

```
proc mixed data=traps;
      class treat block;
      model logcount = treat / ddfm=kr;
      random block;
      * Confidence intervals for treatment means;
      lsmeans treat / cl alpha=0.0125;
      ods output LSMeans=intervals;
run;
data intervals;
      set intervals;
      * Back-transform confidence intervals;
      Mu = 10**(Estimate)-1;
      LowerCL = 10**(lower)-1;
      UpperCL = 10**(upper)-1;
run;
proc print data=intervals;
run;
```

APPENDIX B

Sample SAS program to find confidence intervals for the treatment means using the GLIMMIX macro in SAS (Littell et al., 1996, SAS Institute, 2001). The macro fits a generalized linear mixed model to the data, assuming they are Poisson in distribution and using a log link function to connect the Poisson means to the mixed model. The confidence intervals for the treatment means are generated by the lsmeans statement. One- and two-sided intervals are obtained using the same values of α as in Appendix A. The intervals are automatically back-transformed to the original scale of measurement by the macro.

```
data traps;
      input block treat $ count;
      if treat="BLANK" then delete;
      datalines;
1     AP      4

(data as in Appendix A)

6     ISAP   20
;
run;
%include "glmm800.sas";
%glimmix(data=traps,
      procopt=method=reml,
      stmts=%str(
```

```
        class treat block;
        model count = treat / ddfm=kr;
        random block;
        lsmeans treat / cl alpha=0.0125;),
        error=poisson,
        link=log);
run;
```

Acknowledgments—This study was supported in part by a cooperative agreement with SRS-4501, Southern Research Station, USDA Forest Service. We especially thank two reviewers of the manuscript and the participants of the workshop "Methodological Problems and Solutions in Bark Beetle Semiochemical Research" at the 2003 Southern Forest Insect Work Conference, New Orleans, Louisiana.

REFERENCES

BORDEN, J. H. 1982. Aggregation pheromones, pp. 74–139, *in* J. B. Mitton and K. B. Sturgeon (eds.). Bark Beetles in North American Conifers: A System for the Study of Evolutionary Biology. University of Texas Press, Austin, TX.

BORDEN, J. H. 1997. Disruption of semiochemical-mediated aggregation in bark beetles, pp. 421–438, *in* R. T. Carde and A. K. Minks (eds.). Insect Pheromone Research, New Directions. Kluwer Academic, Boston, MA.

CARMER, S. G. and SWANSON, M. R. 1973. An evaluation of ten pairwise multiple comparison procedures by Monte Carlo methods. *J. Am. Stat. Assoc.* 68:66–74.

CARMER, S. G. and WALKER, W. M. 1982. Baby bear's dilemma: A statistical tale. *Agron. J.* 74:122–124.

COHEN, J. 1988. Statistical Power Analysis for the Behavioral Sciences. Erlbaum, Mahwah, NJ.

DAY, R. W. and QUINN, G. P. 1989. Comparisons of treatments after an analysis of variance in ecology. *Ecol. Monogr.* 59:433–463.

HAYTER, A. J. 1986. The maximum familywise error rate of Fisher's least significant difference test. *J. Am. Stat. Assoc.* 81:1000–1004.

HERMS, D. A., HAACK, R. A., and AYRES, B. D. 1991. Variation in semiochemical-mediated prey–predator interaction: *Ips pini* (Scolytidae) and *Thanasimus dubius* (Cleridae). *J. Chem. Ecol.* 17:1705–1714.

HSU, J. C. 1996. Multiple Comparisons: Theory and Methods. Chapman & Hall/CRC, Boca Raton, Florida.

LITTELL, R. C., MILLIKEN, G. A., STROUP, W. W., and WOLFINGER, R. D. 1996. SAS System for Mixed Models. SAS Institute, Cary, NC.

MCCULLOCH, C. E. and SEARLE, S. R. 2001. Generalized, Linear, and Mixed Models. Wiley, New York, NY.

MILLER, D. R., GIBSON, K. E., RAFFA, K. F., SEYBOLD, S. J., TEALE, S. A., and WOOD, D. L. 1997. Geographic variation in response of pine engraver, *Ips pini*, and associated species to pheromone, lanierone. *J. Chem. Ecol.* 23:2013–2031.

RAFFA, K. F. 2001. Mixed messages across multiple trophic levels: The ecology of bark beetle chemical communication systems. *Chemoecology* 11:49–65.

ROTHMAN, K. J. 1990. No adjustments are needed for multiple comparisons. *Epidemiology* 1:43–46.

SAS INSTITUTE INC. 2001. The SAS System for Windows, Version 8. SAS Institute, Cary, NC.

SAVILLE, D. J. 1990. Multiple comparison procedures: The practical solution. *Am. Stat.* 44:174–180.

SMITH, M. T., SALOM, S. M., and PAYNE, T. L. 1993. The Southern Pine Bark Beetle Guild: An Historical Review of the Research on the Semiochemical-Based Communication System of the Five Principle Species. Virginia Polytechnic Institute and State University, Blacksburg, VA.

SOKAL, R. R. and ROHLF, F. J. 1995. Biometry: The Principles and Practices of Statistics in Biological Research. W. H. Freeman, New York, NY.

STEWART-OATEN, A. 1995. Rules and judgements in statistics: Three examples. *Ecology* 76:2001–2009.

THE R CORE DEVELOPMENT TEAM. 2003. R 1.7.0—A Language and Environment. http://www.r-project.org/

TOOTHAKER, L. E. 1993. Multiple comparison procedures. Sage, Newbury Park, CA.

WESTFALL, P. H., TOBIAS, R. D., ROM, D., WOLFINGER, R. D., and HOCHBERG, Y. 1999. Multiple Comparisons and Multiple Tests Using the SAS System. SAS Institute, Cary, NC.

WOOD, D. L. 1982. The role of pheromones, kairomones, and allomones in the host selection and colonization behavior of bark beetles. *Ann. Rev. Entomol.* 27:411–446.

ZHOU, J., ROSS, D. W., and NIWA, C. G. 2001. Kairomonal response of *Thanasimus undatulus, Enoclerus sphegeus* (Coleoptera: Cleridae), and *Temnochila chlorodia* (Coleoptera: Trogositidae) to bark beetle semiochemicals in Eastern Oregon. *Environ. Entomol.* 30:993–998.